

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. **(currently amended)** A method of evaluating a document set commonality for a document set including a plurality of documents each having one or more document segments, the commonality indicating a degree to which topics of the individual documents of the document set are common, the method comprising:

- (a) generating for each of the document segments, a document segment vector in which components corresponding to terms appearing in the document segment assume a value “1” (one), and the other components assume a value “0” (zero);
- (b) generating for each of the documents of the document set a co-occurrence matrix from the document segment vectors ~~for each of the documents of the document set~~;
- (c) generating a common co-occurrence matrix having rows and columns in which components have values defined as products of values of components of the same rows and the same columns of the [[such]] co-occurrence matrices of the respective documents are given as values of components of the rows and the columns; and
- (d) evaluating the document set commonality on the basis of a sum of either all the components or diagonal components of the common co-occurrence matrix.

2. **(currently amended)** A method of evaluating ~~either~~ a document - document set commonality ~~or a document segment - document set commonality~~ for a document set including a plurality of documents each having one or more document segments, the commonality indicating

a degree to which a document to be evaluated within said document set each of the documents or each of the document segments is close to a topic common to the document set, the method comprising:

- (a) generating for each of the document segments, a document segment vector in which components corresponding to terms appearing in the document segment assume a value “1” (one), and the other components assume a value “0” (zero);
- (b) generating for each of the documents of the document set a co-occurrence matrix from the document segment vectors for each of the documents of the document set;
- (c) generating a common co-occurrence matrix having rows and columns in which components have values defined as products of values of components of the same rows and the same columns of the [[such]] co-occurrence matrices of the respective documents are given as values of components of the rows and the columns; and
- (d) evaluating either the document - document set commonality or the document segment - document set commonality on the basis of
 - either (i) a product-sum between (i1) all the components of either the co-occurrence matrices of matrix for the document to be evaluated and (i2) documents or the document segments and either all the components or diagonal components of the common co-occurrence matrix,
 - or (ii) a product-sum between (ii1) the diagonal components of the co-occurrence matrix for the document to be evaluated and (ii2) the diagonal components of the common co-occurrence matrix.

3. **(currently amended)** A method of calculating a mismatch allowance document set commonality for a document set including a plurality of documents each having one or more document segments, the method comprising:

- (a) generating for each of the document segments, a document segment vector in which components corresponding to terms appearing in the document segment assume a value “1” (one), and the other components assume a value “0” (zero);
- (b) generating for each of the documents of the document set a co-occurrence matrix from the document segment vectors ~~for each of the documents of the document set~~;
- (c) generating a mismatch allowance common co-occurrence matrix having rows and columns ~~in which components have values defined as of a mismatch allowance type from products of non-zero values of components of the same rows and the same columns of the [[such]] co-occurrence matrices of the respective documents, except in cases where the values of the components of the same rows and the same columns are “0” (zero)~~;
- (d) checking if the components of the co-occurrence matrices ~~of the respective documents~~ have the value “0”, and creating a co-occurrence count matrix for counting the number of the documents whose components are not “0”; and
- (e) correcting to become “0” any component of the mismatch allowance common co-occurrence matrix that corresponds to a component when each of components of the co-occurrence count matrix that has a value less than a predetermined threshold, the corresponding component of the common co-occurrence matrix of mismatch allowance type so as to become “0”, and evaluating the mismatch allowance document set commonality of mismatch allowance type on the basis of a sum of either all the components or diagonal components of the corrected mismatch allowance common co-occurrence matrix of a mismatch allowance type.

4. **(currently amended)** A method of extracting documents of common topic from within a document set including a plurality of documents each having one or more document segments, [[;]] the method comprising:

- (a) generating for each of the document segments, a document segment vector in which components corresponding to terms appearing in the document segment assume a value “1” (one), and the other components assume a value “0” (zero);

- (b) generating for each of the documents of the document set a co-occurrence matrix from the [[such]] document segment vectors for each of the documents of the document set;
- (c) generating a mismatch allowance common co-occurrence matrix having rows and columns in which components have values defined as of a mismatch allowance type from products of non-zero values of components of the same rows and the same columns of the [[such]] co-occurrence matrices of the respective documents, except in cases where the values of the components of the same rows and the same columns are “0” (zero);
- (d) checking if the components of the co-occurrence matrices of the respective documents have the value “0”, and creating a co-occurrence count matrix for counting the number of the documents whose components are not “0”;
- (e) correcting to become “0” any component of the mismatch allowance common co-occurrence matrix that corresponds to a component when each of components of the co-occurrence count matrix that has a value less than a predetermined threshold, the corresponding component of the common co-occurrence matrix of a mismatch allowance type so as to become “0”, and evaluating a mismatch allowance document set commonality of a mismatch allowance type on the basis of a sum of either all the components or diagonal components of the corrected mismatch allowance common co-occurrence matrix of mismatch allowance type;
- (f) evaluating a mismatch allowance type document - document set common commonality when the mismatch allowance document set commonality of a mismatch allowance type is not less than a certain threshold, evaluating for each of the documents a mismatch allowance type document - document set common commonality [[and]] on the basis of
 - either (i) a product-sum between (i1) all the components of the co-occurrence matrix [[of]] for the document and (i2) all the components of the corrected mismatch allowance common co-occurrence matrix, of mismatch allowance type
 - or (ii) a product-sum between (ii1) the diagonal components of the co-occurrence matrix [[of]] for the document and (ii2) the diagonal components of

the corrected mismatch allowance common co-occurrence matrix ~~of a mismatch allowance type~~; and

(g) extracting the documents as to which the mismatch allowance type document - document set common commonality exceeds a predetermined threshold, as the documents of a common topic.

5. **(currently amended)** The method as defined in claim 1, further comprising letting M denote the number of sorts of the ~~occurring appearing~~ terms, D_r denote an r^{th} document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y^{th} document segment vector of the document D_r , letter T indicating transposition of a vector, and determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

6. **(currently amended)** The method as defined in claim 2, further comprising letting M denote the number of sorts of the ~~occurring appearing~~ terms, D_r denote an r^{th} document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y^{th} document segment vector of the document D_r , letter T indicating transposition of a vector, and determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

7. **(currently amended)** The method as defined in claim 3, further comprising letting M denote the number of sorts of the ~~occurring appearing~~ terms, D_r denote an r^{th} document in a document set D consisting of R documents, Y_r denote the number of document segments of

the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y th document segment vector of the document D_r , letter T indicating transposition of a vector, and determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

8. **(currently amended)** The method as defined in claim 4, further comprising letting M denote the number of sorts of the ~~occurring appearing~~ terms, D_r denote an r th document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y th document segment vector of the document D_r , letter T indicating transposition of a vector, and determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

9. **(currently amended)** The method as defined in claim [[1]] 5, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

10. **(currently amended)** The method as defined in claim [[2]] 6, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

11. **(currently amended)** The method as defined in claim [[3]] 7, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}.$$

12. **(currently amended)** The method as defined in claim [[4]] 8, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}.$$

13. **(currently amended)** The method as defined in claim 1, further comprising each diagonal component of a common co-occurrence matrix S^C of a document set D being approximated by a product of occurring frequencies of [[the]] each corresponding term in the respective documents.

14. **(currently amended)** The method as defined in claim 2, further comprising each diagonal component of a common co-occurrence matrix S^C of a document set D being approximated by a product of occurring frequencies of [[the]] each corresponding term in the respective documents.

15. **(currently amended)** The method as defined in claim 3, further comprising each diagonal component of a common co-occurrence matrix S^C of a document set D being approximated by a product of occurring frequencies of [[the]] each corresponding term in the respective documents.

16. **(currently amended)** The method as defined in claim 4, further comprising each diagonal component of a common co-occurrence matrix S^C of a document set D being approximated by a product of occurring frequencies of [[the]] each corresponding term in the respective documents.

17. (original) A program storage device, readable by a machine, tangibly embodying a program of instructions executable by the machine to perform the method of claim 1.

18. (original) A program storage device, readable by a machine, tangibly embodying a program of instructions executable by the machine to perform the method of claim 2.

19. (original) A program storage device, readable by a machine, tangibly embodying a program of instructions executable by the machine to perform the method of claim 3.

20. (original) A program storage device, readable by a machine, tangibly embodying a program of instructions executable by the machine to perform the method of claim 4.

21. **(currently amended)** The program storage device as defined in claim 17, further comprising letting M denote the number of sorts of the occurring appearing terms, D_r denote an r^{th} document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y^{th} document segment vector of

the document D_r , letter T indicating transposition of a vector, determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

22. **(currently amended)** The program storage device as defined in claim 18, further comprising letting M denote the number of sorts of the occurring appearing terms, D_r denote an r th document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y th document segment vector of the document D_r , letter T indicating transposition of a vector, determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

23. **(currently amended)** The program storage device as defined in claim 19, further comprising letting M denote the number of sorts of the occurring appearing terms, D_r denote an r th document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y th document segment vector of the document D_r , letter T indicating transposition of a vector, determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

24. **(currently amended)** The program storage device as defined in claim 20, further comprising letting M denote the number of sorts of the occurring appearing terms, D_r denote an r th document in a document set D consisting of R documents, Y_r denote the number of document segments of the document D_r , and $d_{ry} = (d_{ry1}, \dots, d_{ryM})^T$ denote a y th document segment vector of

the document D_r , letter T indicating transposition of a vector, determining the co-occurrence matrix S^r of the document D_r by:

$$S^r = \sum_{y=1}^{Y_r} d_{ry} d_{ry}^T$$

25. **(currently amended)** The program storage device as defined in claim [[17]] 21, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

26. **(currently amended)** The program storage device as defined in claim [[18]] 22, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

27. **(currently amended)** The program storage device as defined in claim [[19]] 23, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

28. **(currently amended)** The program storage device as defined in claim [[20]] 24, further comprising determining an mn component S^C_{mn} of a common co-occurrence matrix S^C of a document set D by:

$$S^C_{mn} = \prod_{r=1}^R S^r_{mn}$$

29. (original) A computer system arranged to perform the method of claim 1.

30. (original) A computer system arranged to perform the method of claim 2.

31. (original) A computer system arranged to perform the method of claim 3.

32. (original) A computer system arranged to perform the method of claim 4.

33. (**new**) A method of evaluating a document segment - document set commonality for a document set including a plurality of documents each having one or more document segments, the commonality indicating a degree to which a document segment to be evaluated in a document within said document set is close to a topic common to the document set, the method comprising:

(a) generating for each of the document segments, a document segment vector in which components corresponding to terms appearing in the document segment assume a value “1” (one), and the other components assume a value “0” (zero);

(b) generating for each of the documents of the document set a co-occurrence matrix from the document segment vectors;

(c) generating a common co-occurrence matrix having rows and columns in which components have values defined as products of values of components of the same rows and the same columns of the co-occurrence matrices;

(d) evaluating a document set commonality, that indicates a degree to which topics of the individual documents of the document set are common, on the basis of a sum of either all the components or diagonal components of the common co-occurrence matrix; and

(e) evaluating the document segment - document set commonality on the basis of the common co-occurrence matrix.